

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using SW model

Run on: January 29, 2004, 23:04:15 ; Search time 40 Seconds
(without alignments)
1190.449 Million cell updates/sec

Title: US-09-820-598-1

Perfect score: 1589
Sequence: 1 MHKMEKSONFRLEALAEK.....DEBEDDQMDTAERKMGSVL 300

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: A.GeneSeq_19Jun03.*
2: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1980.DAT.*
3: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1981.DAT.*
4: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1982.DAT.*
5: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1983.DAT.*
6: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1984.DAT.*
7: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1985.DAT.*
8: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1589	100.0	300	21	AAV84616
2	620.5	39.0	349	21	AAV84617
3	376	23.7	525	22	ABB63763
4	270.5	17.0	494	22	ABB59553
5	266.5	16.8	283	20	AAW73583
6	266.5	16.8	283	22	AAE0726
7	266.5	16.8	283	24	ABP73014
8	266.5	16.8	416	23	AAE35660
9	266.5	16.8	416	24	ABP97648

10	266	16.7	284	21	AA26767	Rat insulin promot
11	266	16.7	284	22	AA35426	Secretory cell lin
12	266	16.7	284	24	ABP5166	Mouse transcriptio
13	265.5	16.7	283	16	AA70031	Insulin transcript
14	263	16.6	524	22	ABB64765	Drosophila melanog
15	259	16.3	356	24	ABR47489	Breast cancer asse
16	254	16.0	284	20	AAW83396	Mouse insulin prom
17	254	16.0	284	23	AAU99330	Mouse insulin prom
18	254	16.0	284	23	AAU99330	Mouse insulin prom
19	250.5	15.8	301	20	AAW94463	Human wild-type Ho
20	250.5	15.8	304	20	AAW94464	Human HoxB1 varian
21	248	15.6	376	23	ABP62865	Human polypeptide
22	247	15.5	485	22	ABB58326	Drosophila melanog
23	245	15.4	777	22	ABB66529	Drosophila melanog
24	243	15.3	252	23	ABP68997	Human polypeptide
25	243	15.3	252	23	AAE25284	Human nucleic acid
26	243	15.3	252	23	ABR05676	Human transcriptio
27	242.5	15.3	367	22	ABG26521	Novel human diagno
28	242.5	15.3	465	22	ABB57996	Drosophila melanog
29	242	15.2	368	23	ABP69498	Human polypeptide
30	238	15.0	260	23	ABP43887	Human protein cont
31	236	14.9	220	22	AAW41487	Human polypeptide
32	235.5	14.8	299	23	ABR08329	Mouse homeo box ms
33	235.5	14.8	378	22	ABR72035	Drosophila melanog
34	234	14.7	304	24	AAO16404	Human nucleic acid
35	233	14.7	357	22	ABR10101	Human homeobox pro
36	232	14.6	335	20	AAW94465	Human HoxA1 varian
37	232	14.6	335	20	AAW94462	Human wild-type Ho
38	231.5	14.6	391	22	ABR62664	Drosophila melanog
39	230.5	14.5	236	22	AAW39701	Human polypeptide
40	229.5	14.4	431	22	AAW95273	Human protein sequ
41	228.5	14.4	302	16	AAW82097	Human Gax protein.
42	225	14.2	230	21	AAW58610	Protein regulating
43	224.5	14.1	271	22	ABG18979	Novel human diagno
44	224	14.1	367	23	ABG70892	Human NKX-6.1 proc
45	221	13.9	217	23	ABG96337	Human ovarian Canc

ALIGNMENTS

RESULT 1	AAV84616	standard; Protein; 300 AA.
XX	AAV84616;	
AC	AAV84616;	
XX	25-JUL-2000 (first entry)	
DT	25-JUL-2000 (first entry)	
XX		
DE	A chicken motor neuron restricted pattern 2 protein.	
XX		
KM	Motor neuron restricted pattern 2 protein; MNR2 protein; motor neuron;	
KM	Pax6+ motor neuron progenitor; somatic motor neuron;	
KM	transcription factor; Isl2; Lim 3; homeobox protein 9; HB9;	
KM	chronic neurodegenerative disease; spinal muscular atrophy;	
KM	myotrophic lateral sclerosis; neuromuscular disease; sacral agenesis.	
XX		
OS	Gallus sp.	
XX		
PN	MO200018884-A1.	
PD	06-APR-2000.	
XX		
PF	29-SEP-1999; 99MO-US22517.	
XX		
PR	29-SEP-1998; 98US-0162524.	
XX		
PA	(UYCO) UNIV COLUMBIA NEW YORK.	
XX		
PI	Jeeseil TM, Tanabe Y, William C;	
XX		
XX	WPI; 2000-293133/25.	
DR	N-PSDB; AA299988.	

XX New isolated nucleic acid molecules encoding motor neuron restricted
 PT pattern protein and a homeobox protein are used in treating chronic
 PT neurodegenerative diseases and acute nervous system injuries -
 XX
 XX
 PS Claim 31; Fig 10; 209pp; English.

CC The present sequence represents a motor neuron restricted pattern (MNR2)
 CC protein from a chicken embryo. MNR2 is expressed selectively by Pax6+
 CC motor neuron progenitors and persists in post-mitotic somatic motor
 CC neurons. Expression of the MNR2 protein induces expression of
 CC transcription factors Isl2, Lim 3 and homeobox protein 9 (HB9).
 CC MNR2 and HB9 are used to induce differentiation of somatic motor
 CC neurons in a mammal which is to treat an abnormality associated with a
 CC lack of one or more normally functioning motor neurons, a chronic
 CC neurodegenerative disease such as spinal muscular atrophy or myotrophic
 CC lateral sclerosis, an acute nervous system injury localized to a
 CC specific central axon and neuromuscular disease. HB9 is also used to
 CC treat an embryo afflicted with sacral agenesis.

XX Sequence 300 AA;

Query Match 100.0%; Score 1589; DB 21; Length 300;
 Best Local Similarity 100.0%; Pred. No. 1.8e-125;
 Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEKPEKSONFRITELALAEKPPRSASPGGLSPAGSGPGAGRTDTPSPAPQATPLGAG 60
 DB 1 MEKPEKSONFRITELALAEKPPRSASPGGLSPAGSGPGAGRTDTPSPAPQATPLGAG 60
 QY 61 FVPRKGLHLTPGPGGLTLPALYPAVYPLPALGQHAAPAYTAPFOLPPPGAHLKAAYA 120
 DB 61 FVPRKGLHLTPGPGGLTLPALYPAVYPLPALGQHAAPAYTAPFOLPPPGAHLKAAYA 120
 QY 121 GSPFLEQWIRAGMLYPRUSDFAHATQOSALMGSKSRPRPTFTSQQLLEENOKLANKYLSR 180
 DB 121 GSPFLEQWIRAGMLYPRUSDFAHATQOSALMGSKSRPRPTFTSQQLLEENOKLANKYLSR 180
 QY 181 PRFPAVATSLMTEOVKIMFONRRMKRSRKAKQGMVAPEKPRGLGKADESLPSQ 240
 DB 181 PRFPAVATSLMTEOVKIMFONRRMKRSRKAKQGMVAPEKPRGLGKADESLPSQ 240
 QY 241 PGGQAGDSPEFVGCSPGTGFLCRSAELGYDPDSCSGGEDEEBEDDGMTAERMGSYL 300
 DB 241 PGGQAGDSPEFVGCSPGTGFLCRSAELGYDPDSCSGGEDEEBEDDGMTAERMGSYL 300

RESULT 2
 AAY84617
 ID AAY84617 standard; Protein; 349 AA.

AC AAY84617;

DT 25-JUL-2000 (first entry)

DE Amino acid sequence of the chicken homeobox protein 9 (HB9).

XX Motor neuron restricted pattern 2 protein; MNR2 protein; motor neuron;
 XX Pax6+ motor neuron progenitor; somatic motor neuron;
 XX transcription factor; Isl2; Lim 3; homeobox protein 9; HB9;
 XX chronic neurodegenerative disease; spinal muscular atrophy;
 XX myotrophic lateral sclerosis; neuromuscular disease; sacral agenesis.

OS Gallus sp.

PN WO200018884-A1.

PD 06-APR-2000.

PF 29-SEP-1999; 99WO-US22517.

PR 29-SEP-1998; 98US-0162524.

XX

PA (UYCO) UNIV COLUMBIA NEW YORK.

PI Jessell TM, Tanabe Y, William C;

XX WPI; 2000-29133/25.

DR N-PSDB; AA299989.

PT New isolated nucleic acid molecules encoding motor neuron restricted
 PT pattern protein and a homeobox protein are used in treating chronic
 PT neurodegenerative diseases and acute nervous system injuries -
 XX
 XX
 PS Disclosure; Fig 12; 209pp; English.

CC The present sequence represents a homeobox protein 9 (HB9). The
 CC specification describes a motor neuron restricted pattern (MNR2)
 CC protein from a chicken embryo. MNR2 is expressed selectively by Pax6+
 CC motor neuron progenitors and persists in post-mitotic somatic motor
 CC neurons. Expression of the MNR2 protein induces expression of
 CC transcription factors Isl2, Lim 3 and homeobox protein 9 (HB9).
 CC MNR2 and HB9 are used to induce differentiation of somatic motor
 CC neurons in a mammal which is to treat an abnormality associated with a
 CC lack of one or more normally functioning motor neurons, a chronic
 CC neurodegenerative disease such as spinal muscular atrophy or myotrophic
 CC lateral sclerosis, an acute nervous system injury localized to a
 CC specific central axon and neuromuscular disease. HB9 is also used to
 CC treat an embryo afflicted with sacral agenesis.

XX Sequence 349 AA;

Query Match 39.0%; Score 620.5; DB 21; Length 349;
 Best Local Similarity 47.3%; Pred. No. 5.5e-44;
 Matches 159; Conservative 30; Mismatches 68; Indels 79; Gaps 14;

QY 5 MEKSONFRITELALAEKPPRSASPGGLSPAGSGPGAGRTDTPSPAPQATPLGAG 39
 DB 1 MEKSONFRITELALAEKPPRSASPGGLSPAGSGPGAGRTDTPSPAPQATPLGAG 39
 QY 40 GRTDTPSP-----RAPQA--TFLGPAAGFVPRKGL--LHLPGGLTLPALYPAV 86
 DB 61 PRDTPSPRLLPALHCAFLPRALVLTGSGGSGSPSSSSSSSSSSSELPLADC 60
 QY 87 YPLPALGQHAAPAYT-----AFPOLPPGAHLKAAYVGSFLEQWIR---AGMLVPR 137
 DB 118 YGYPALGQHPALSYSSVOGAPR--AHPSADPIKLS--AGTFLDQMLASTAGMTLPK 174
 QY 138 LSDFAHATQOSALMGSKSRPRPTFTSQQLLEENOKLANKYLSRKPREVATSLMTEOV 197
 DB 175 MPDQSQAGSNLTKCRPRPTFTSQQLLEENOKLANKYLSRKPREVATSLMTEOV 234
 QY 198 KIMFONRRMKRSRKAKQGMVAPEKPRGLGKAD-----ESLPSQPGQAGDSPEF 251
 DB 235 KIMFONRRMKRSRKAKQGMVAPEKPRGLGKAD-----ESLPSQPGQAGDSPEF 288
 QY 252 VGCSPGTGFLCRSAELGYDPDSCSGGEDEEBEDD 287
 DB 289 -----RLREL--PDSEPEDEEBEEDDEE 310

RESULT 3
 ABB63763
 ID ABB63763 standard; Protein; 525 AA.

AC ABB63763;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 18081.

KM Drosophila; developmental biology; cell signalling; insecticide;
 KM pharmaceutical.

OS Drosophila melanogaster.

XX

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 23:06:10 / Search time 20 Seconds

(without alignments)
1442.530 Million cell updates/sec

Title: US-09-820-598-1

Perfect score: 1589

Sequence: 1 MHKMEKSNQFRLEALLAEK.....DEEEDDDGMTAKKMSVL 300

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	583	36.7	403	2 A53662	homeotic protein H
2	413	26.0	316	2 S16681	homeotic protein -
3	288	18.1	309	1 S09257	homeotic protein H
4	285.5	18.0	261	2 S63604	homeobox protein G
5	284	17.9	285	1 A43556	homeotic protein H
6	282	17.7	255	1 M3HUB	homeotic protein H
7	274	17.2	283	2 G01926	insulin promoter f
8	270.5	17.0	497	2 S22708	homeotic protein e
9	267.5	16.8	250	1 A36170	homeotic protein H
10	266.5	16.8	283	2 A42634	homeoprotein, IDX-
11	266.5	16.8	416	2 S27198	homeotic protein H
12	266	16.7	284	2 S39581	IPFI protein - mou
13	264.5	16.6	375	2 S50589	transcription fact
14	264.5	16.6	375	2 S35941	homeotic protein H
15	261	16.4	320	1 A39724	homeotic protein H
16	261	16.4	417	2 S47539	homeotic protein H
17	260.5	16.4	411	2 S34164	homeotic protein H
18	259	16.3	356	1 M3HUB	homeotic protein a
19	255.5	16.1	295	2 S58850	homeotic protein H
20	255.5	16.1	348	2 S52720	homeobox protein g
21	252	15.9	243	2 JC7556	GXI protein - chl
22	250.5	15.8	284	2 B41224	homeotic protein P
23	248.5	15.6	301	1 M3HUB	homeotic protein H
24	248.5	15.6	245	2 T46446	hypothetical prote
25	248.5	15.6	251	1 B60492	homeotic protein H
26	248	15.6	250	1 A31757	homeotic protein H
27	245.5	15.4	277	2 S78063	homeobox protein P
28	245.5	15.4	798	2 S20881	homeotic protein P
29	245	15.4	264	1 S35219	homeotic protein H

30	245	15.4	372	2 A46037	Hox-1.11 - mouse
31	244	15.4	372	2 I52196	homeobox transcrip
32	243.5	15.3	297	2 A40560	homeotic protein H
33	243.5	15.3	327	2 A56553	homeotic protein H
34	242.5	15.3	297	1 M3MS29	homeotic protein H
35	241.5	15.2	297	2 I54320	homeobox protein -
36	240	15.1	288	2 US0659	homeotic protein H
37	239	15.0	270	1 M3MS13	homeotic protein H
38	239	15.0	309	1 A60096	homeotic protein H
39	238	15.0	326	1 S09256	homeotic protein H
40	238	15.0	260	2 A43821	homeotic protein A
41	237.5	14.9	394	2 S26492	homeotic protein A
42	236.5	14.9	271	2 S30230	homeotic protein H
43	235.5	14.8	299	2 S18813	homeotic protein H
44	235.5	14.8	323	2 A54772	homeotic protein M
45	235.5	14.8	378	2 A25399	homeotic protein A

ALIGNMENTS

RESULT 1

A53662
homeotic protein HB9 - human
C/Species: Homo sapiens (man)
C/Date: 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 17-Oct-1997
C/Accession: A53662
R/Harrison, K.A.; Druey, K.M.; Deguchi, Y.; Tusciano, J.M.; Kehrl, J.H.
J. Biol. Chem. 269, 19968-19975, 1994
A/Title: A novel human homeobox gene distantly related to proboscipedia is expressed in
A/Reference number: A53662; MUID:94327547; PMID:7914194
A/Accession: A53662
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: DNA
A/Residues: 1403 <HAR>
A/Cross-references: GB:007663
A/Note: the nucleotide sequence and conceptual translation as given are self-consistent
A/Genes:
A/Gene: GDB:HLX89
A/Cross-references: GDB:136411; OMIM:142994
A/Map position: 1q41-q42.1
A/Superfamily: unassigned homeobox proteins; homeobox homology
C/Keywords: DNA binding; homeobox; nucleus; transcription regulation
F/244-300/Domain: homeobox homology <Hox>

Query Match 36.7%; Score 583; DB 2; Length 403;
Best Local Similarity 41.1%; Pred. No. 4.4e-30;
Matches 162; Conservative 30; Mismatches 64; Indels 138; Gaps 16;

QY	5	MEKSNQFRLEALLAEKPPRSAS	-----PPGL-----	SP 32
DB	1	MEKSNQFRLEALLAEKPPRSAS	-----PPGL-----	60
QY	33	AGSPGPGAG-----RTDTPSPRAPOAATPLGPAGFVFKPGILHLPGLGT	-----	77
DB	61	ASEPRAAPADRLRAESPSPRLAA-----HCALLPKPGFLGAGGGCGGCGGHPHH	-----	116
QY	78	-----LP-----ALYPAVY-----PLPA 91		
DB	117	AHFGAAAAAAGGALGLHPGAGAGGLPQAALYGHPRVGYSAAAAAA 176		
QY	92	LGGOHAFAVTAFFPOL-----PPGAEHLKAAVAGSPFLEQWR-----AGMLVPLSPFHAT 144		
DB	177	LACQHPALSY-STPVOQAHPAHPADPILKG-AGTQLODMILPASTAGMLPMPDPNSQ 234		
QY	145	POGALMGKSRPRTAFTSOQLLENOFKLNTKLSRPKREVAVSLMTETQYKIFQNR 204		
DB	235	AGSNLKGCRPRPTAFTSOQLLENOFKLNTKLSRPKREVAVSLMTETQYKIFQNR 294		
QY	205	RMKTKSRKAKEGMAVEBEKPGLG-----KADESLLPQPGQADSPFVG 253		
DB	295	RMKTKSRKAKEGMAVEBEKPGLG-----KADESLLPQPGQADSPFVG 344		

Qy 254 CSPGTGFLCRSAELGYDPDSSCGGEDEDEEDD 287
 Db 345 ---GSGRRLR-----DLRDSPEDEDEDE 367

RESULT 2

homeotic protein - human
 C/Species: Homo sapiens (man)
 C/Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 17-Oct-1997
 C/Accession: S16681
 R/Deguchi, Y.; Kehr1, J.H.
 Nucleic Acids Res. 19, 3742, 1991
 A/Title: Nucleotide sequence of a novel diverged human homeobox gene encodes a DNA bindi
 A/Reference number: S16681; MUID:91305125; PMID:1677181
 A/Accession: S16681
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: mRNA
 A/Residues: 1-316 <DEG>
 A/Cross-references: EMBL:X56537
 A/Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1991
 C/Superfamily: unassigned homeobox proteins; homeobox homology
 C/Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F/73-127/Domain: homeobox homology <HOX>

Query Match 26.0%; Score 413; DB 2; Length 316;
 Best Local Similarity 48.8%; Pred. No. 2.2e-19;
 Matches 106; Conservative 24; Mismatches 41; Indels 46; Gaps 11;

Qy 83 PPAYPLPALGGQHAFAFYTAFPOLPPGAELKAANAAGSFPLEQWIR--AGMLVPLRLS 139
 Db 11 PGALLLVPA-GARRA-----PRDPADPIKLG-AGTFQLDQWLRASTAGMILPMP 58
 Qy 140 DFHATPSALMGKRRRTAFTSQQLLENOFLKNTLSRPKEVATSLMETQYKI 199
 Db 59 DFNSQASNLGKCRSRSTAFTSQQLLEL--FFKNKYLSPKKEVATSLMETQYKI 116
 Qy 200 WFORRMKMKRSRAKEQGMAVEPEK--PRGLGK-----ADSLSPSQGQAGSPPE 250
 Db 117 WFORRMKMKRSKKAKEKRRKRRNRRAAGPCKFGAEPEFAEELGPAPF----GDK-- 170
 Qy 251 FVGCSPGTGFLCRSAELGYDPDSSCGGEDEDEEDD 287
 Db 171 ---GSGRRLRRLR---DSDP-----EEDDEDDE 193

RESULT 3

homeotic protein Hox A4 - chicken
 N/Alternate names: homeotic protein Chox1-4
 C/Species: Gallus gallus (chicken)
 C/Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 21-Jul-2000
 C/Accession: S09257; S10883
 R/Sasaki, H.; Yokoyama, E.; Kuroiwa, A.
 Nucleic Acids Res. 18, 1739-1747, 1990
 A/Title: Specific DNA binding of the two chicken deformed family homeodomain proteins,
 A/Reference number: S09256; MUID:90245562; PMID:1970866
 A/Accession: S09257
 A/Molecule type: mRNA
 A/Residues: 1-309 <SAS>
 A/Cross-references: EMBL:X52670; NID:G63218; PIDN:CNA36896.1; PID:G63219
 R/Scottling, P.J.; Hewitt, M.; Keynes, R.J.
 Nucleic Acids Res. 18, 3999, 1990
 A/Title: Isolation and analysis of chick homeobox cDNA clones.
 A/Reference number: S10883; MUID:90326535; PMID:1973835
 A/Accession: S10883
 A/Molecule type: mRNA
 A/Residues: 207-273 <SCO>
 A/Cross-references: EMBL:X52747; NID:G63223; PIDN:CBM57949.1; PID:G6018426
 C/Function:
 A/Description: control of embryonic development by tissue- and stage-specific regulation
 C/Superfamily: homeotic protein Hox D4; homeobox homology
 C/Keywords: DNA binding; embryo; homeobox; nucleus; transcription regulation

F/210-266/Domain: homeobox homology <HOX>

Query Match 18.1%; Score 288; DB 1; Length 309;
 Best Local Similarity 32.2%; Pred. No. 1.8e-11;
 Matches 88; Conservative 27; Mismatches 88; Indels 70; Gaps 10;

Qy 18 AEKPRRSASPPGLSPAGSPGAGRTDTPSP-RAPOAATPLGPAQFVFKPGLHLPGGLG 76
 Db 45 APPPPPPPPPHLH-AAHPGALBEYFPPRRRERGVQAPAAP-----PGP--- 88
 Qy 77 TLPAIYPPAYPLPALGGQHAFAFYTAFPOLPPGA-----POLPPGA----- 112
 Db 89 --PQPPPEALYPPQAPSPYQAPYSPYSAGSAGAPQEPQPPGASPPPPPAKGGHPAQL 146
 Qy 113 -----EHLKAANAAGSF--LEQWIRAGM-----LVPRLSDFATPQSALM-- 150
 Db 147 LPHGALORCEAPAPAAAGAGPGGALLPDXSLPDLKKEEVVYPMWKIIVSTNPNYSG 206
 Qy 151 GKSRPRRTAFTSQQLLENOFLKNTLSRPKEVATSLMETQYKIFONRMKMKR 210
 Db 207 GEPKRSRTATYTRQVLELEKEFFHNRRLTRRRRIEIAHTLCLSERQYKIFONRMKMKK 266
 Qy 211 SRKAKEQMAVEPEKPRGLKADBSLLPSQPG 243
 Db 267 DHKLPNTKMRSSNQPSLGGQQA-----KAOTOG 294

RESULT 4

homeobox protein Gsh-1 - mouse
 C/Species: Mus musculus (house mouse)
 C/Date: 28-Oct-1996 #sequence_revision 27-Feb-1997 #text_change 17-Nov-2000
 C/Accession: S63604; S66126; A37290; A38809
 R/Li, H.; Zeller, P.S.; Valerius, M.T.; Small, K.; Potter, S.S.
 EMBO J. 15, 714-724, 1996
 A/Title: Gsh-1, an orphan hox gene, is required for normal pituitary development.
 A/Reference number: S63604; MUID:96181350; PMID:8631293
 A/Accession: S63604
 A/Status: not compared with conceptual translation
 A/Molecule type: DNA
 A/Residues: 1-261 <LTA>
 R/Valerius, M.T.; Li, H.; Stock, J.L.; Weinstein, M.; Kaur, S.; Singh, G.; Potter, S.S.
 Dev. Dyn. 203, 337-351, 1995
 A/Title: Gsh-1: A novel murine homeobox gene expressed in the central nervous system.
 A/Reference number: S66126; MUID:9612995; PMID:8589431
 A/Accession: S66126
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-261 <VAL>
 A/Cross-references: EMBL:U21224; NID:G836957; PIDN:AAA96814.1; PID:G836958
 R/Singh, G.; Kaur, S.; Stock, J.L.; Jenkins, N.A.; Gilbert, D.J.; Copeland, N.G.; Potter,
 Proc. Natl. Acad. Sci. U.S.A. 88, 10706-10710, 1991
 A/Title: Identification of 10 murine homeobox genes.
 A/Reference number: A37290; MUID:92073356; PMID:1683797
 A/Accession: A37290
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 146-205 <SIN>
 C/Superfamily: unassigned homeobox proteins; homeobox homology
 C/Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F/147-203/Domain: homeobox homology <HOX>

Query Match 18.0%; Score 285.5; DB 2; Length 261;
 Best Local Similarity 33.8%; Pred. No. 2.2e-11;
 Matches 100; Conservative 28; Mismatches 105; Indels 63; Gaps 13;

Qy 9 QNFRLEALAEKPPRSAPGLSPAGSPGAGRTDTPSPAPQAATPLGPAQFVFKPGL 68
 Db 3 RSLFVDSLVL---REASDK-KAPBGSPPPLFYAVNPFRHALLGLSP--GACARAKGLL 55
 Qy 69 HL-----PG-PGLGTLPALYPP--AVYPLPALGGQHAFAFYTAFPOLPPGA 112
 Db 56 CVCPICVTASQLHGPPGPPALPLTKASFPPFGSQYCHAPLGRHSV-----SPGV 105

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using SW model

Run on: January 29, 2004, 23:05:05 ; Search time 17 seconds

(without alignments)
829.883 Million cell updates/sec

Title: US-09-820-598-1

Sequence: 1 MHKPEKSNQFRIFALLAEK.....DEEEDDDGMDTAEKMGSVL 300

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt_41*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	561	35.3	401	HB9_HUMAN	P50219 homo sapien
2	299.5	18.8	340	GBX2_CHICK	O42230 gallus gall
3	288	18.1	364	GSHI_HUMAN	Q9H482 homo sapien
4	288	18.1	309	HXA4_CHICK	P17277 gallus gall
5	285.5	18.0	261	GSHI_MOUSE	P13315 mus musculu
6	284	17.9	285	HXA4_MOUSE	P06798 mus musculu
7	274	17.2	255	HXD4_HUMAN	P09016 homo sapien
8	270.5	17.0	283	IPF1_HUMAN	P52945 h insulin p
9	270.5	17.0	497	HME5_DROME	P18488 drosophila
10	269.5	17.0	443	HXA3_MOUSE	P02831 mus musculu
11	268.5	16.9	283	IPF1_MOUSE	P70118 mesocricetu
12	267.5	16.8	250	HXD4_MOUSE	P10628 mus musculu
13	266.5	16.8	416	HXD3_HUMAN	P51249 homo sapien
14	266	16.7	284	IPF1_MOUSE	P52947 mus musculu
15	265.5	16.7	283	IPF1_RAT	P52947 rattus norv
16	265	16.7	443	HXA3_HUMAN	O43365 homo sapien
17	264.5	16.6	254	BRX1_MOUSE	O98742 mus musculu
18	264.5	16.6	375	HXA2_CHICK	Q08727 gallus gall
19	261	16.4	417	HXD3_MOUSE	P09027 mus musculu
20	259	16.3	356	HXB2_HUMAN	P16522 homo sapien
21	258	16.2	297	TLX3_CHICK	O93367 gallus gall
22	257	16.2	320	HXA4_HUMAN	O00056 homo sapien
23	255.5	16.1	207	BRX1_CHICK	O96688 gallus gall
24	255.5	16.1	348	GBX2_MOUSE	P46031 mus musculu
25	253.5	16.0	413	HXD3_CHICK	O93353 gallus gall
26	252.5	15.9	348	GBX2_HUMAN	P52951 homo sapien
27	250.5	15.8	284	TLX2_MOUSE	O61663 mus musculu
28	250.5	15.8	301	HXB1_HUMAN	P14653 homo sapien
29	248.5	15.6	251	HXB4_HUMAN	P17483 homo sapien
30	248	15.6	250	HXA2_MOUSE	P10282 mus musculu
31	248	15.6	376	HXA2_HUMAN	O43364 homo sapien
32	247	15.5	271	HME5_XENLA	P14837 xenopus lae
33	246.5	15.5	275	HXA5_HETFR	O91a23 heterodontu

34	245.5	15.4	277	1	HMP1_CHICK	O05502 gallus gall
35	245.5	15.4	788	1	HMP1_DROME	P31264 drosophila
36	245	15.4	264	1	HXC4_MOUSE	O08624 mus musculu
37	245	15.4	372	1	HXA2_MOUSE	P31245 mus musculu
38	244	15.4	372	1	HXA2_RAT	P31246 rattus norv
39	243.5	15.3	327	1	HXD1_MOUSE	O01822 mus musculu
40	243	15.3	252	1	HXA2_HUMAN	O04743 homo sapien
41	242.5	15.3	297	1	HXB1_MOUSE	P17919 mus musculu
42	242.5	15.3	515	1	HMSH_DROME	O03372 drosophila
43	242	15.2	328	1	HXD1_HUMAN	O99220 homo sapien
44	241.5	15.2	297	1	MSX1_HUMAN	P28360 homo sapien
45	241	15.2	363	1	HXA2_HETFR	O91a20 heterodontu

ALIGNMENTS

RESULT 1
HB9_HUMAN STANDARD; PRT; 401 AA.
ID HB9_HUMAN
AC P50219;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Homeobox protein HB9.
GN HLBX9.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=94327547; Pubmed=7914194;
RA Harrison K.A., Druey K.M., Deguchi Y., Tusciano J.M., Kehrl J.H.;
RT "A novel human homeobox gene distantly related to proboscipedia is
RL J. Biol. Chem. 269:19968-19975(1994).
CC -1- FUNCTION: PUTATIVE TRANSCRIPTION FACTOR.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN LYMPHOID AND PANCREATIC TISSUES.
CC -1- SIMILARITY: Contains 1 Homeobox domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: U07664; AAB6047.1; -
DR EMBL: U07663; AAB6047.1; JOINED.
DR HSSP: P14653; 1872.
DR TRANSFAC: T03420; -
DR Genew: HGNC:4979; HLBX9.
DR MIT: 142994; -
DR GO: GO:0005634; C:nucleus; NAS.
DR GO: GO:0003702; F:RNA polymerase II transcription factor acti. .; TAS.
DR GO: GO:0003700; F:transcription factor activity; NAS.
DR GO: GO:0007345; P:embryogenesis and morphogenesis; TAS.
DR GO: GO:0006959; P:numeral immune response; TAS.
DR GO: GO:0006357; P:regulation of transcription from Pol II pro. .; TAS.
DR InterPro: IPR001356; Homeobox.
DR Pfam: PF00046; homeobox; 1.
DR ProDom: PD000010; Homeobox; 1.
DR SMART: SM00389; HOX; 1.
DR PROSITE: PS00027; HOMEBOX_1; 1.
DR PROSITE: PS50071; HOMEBOX_2; 1.
DR Homeobox; DNA-binding; Nuclear protein; Transcription regulation.
FT DOMAIN 39 48 POLY-GLY.
FT DOMAIN 97 111 POLY-GLY.
FT DOMAIN 120 135 POLY-ALA.

FT DOMAIN 169 177 POLY-ALA.
 FT DNA BIND 242 HOMEBOX.
 FT DOMAIN 316 325 POLY-GLY.
 SQ SEQUENCE 401 AA; 40932 MW; 0006AED71D594FE CRC64;

Query Match 35.3%; Score 561; DB 1; Length 401;
 Best Local Similarity 39.7%; Pred. No. 1.7e-26;
 Matches 156; Conservative 30; Mismatches 69; Indels 138; Gaps 15;

QY 5 MEKSONFRIEAL-----IAEK-----PPSASPPG-----LSPA 33
 DB 1 MEKSNFRIEPCWMTPEHPLEALAKVTSPPVPASGTGGGGGSGSGSCSPA 60
 QY 34 GSPGPAQ-----RTDTPSPAPQAATPLGPAQFVFKPGLHLPGLGT-----77
 DB 61 SSEPPLAADRLRAESPPRLLAA-----HCALLPKPFLGAGGGGGTGGHGPNNHA 116
 QY 78 -----LP-----ALYPVAVY-----PLPAL 92
 DB 117 HPGAAAAAAGGALALGHPGAAGAGLPAQALYGHVGYSAAMAAAL 176
 QY 93 GGGHAAFAVTAEPOL-----PPGAHLLKAAVAGSPLEQWTR--AGMLVRLSDFAHP 145
 DB 177 AGQHPALSY-SYPOVQGAHPAPADPIKLG-AGTFQLDQMLRASTGMILKMPDPSQA 234
 QY 146 OSALMGKRRPRTAFTSQQLLENOFLNKYLSRPRFEVATSLMTETQVKIWFQWR 205
 DB 225 OSNLGKRRRTAFTSQQLLENOFLNKYLSRPRFEVATSLMTETQVKIWFQWR 294
 QY 206 MKMKRSRAKQGMVAVEPEKPGILG-----KADESILPSQPOGAQDSPEFVGC 254
 DB 295 MKMKRSRAKQGMVAVEPEKPGILG-----KADESILPSQPOGAQDSPEFVGC 348
 QY 255 SPGTGFLCRSAELGDPSSCGSGEDEDDEED 287
 DB 349 -----ADLRSDPEDEDEDE 365

RESULT 2

GBX2_CHICK STANDARD; PRT; 340 AA.
 AC 042230;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Homeobox protein GBX-2 (Gastrulation and brain-specific homeobox protein 2).
 GN GBX2.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
 OX NCBI_TaxID=9031;
 RN 11
 RP MEDLINE=98004222; PubMed=9346236;
 RA Kowenz-Leutz E., Herr P., Niss K., Leutz A.;
 RT "The homeobox gene GBX2, a target of the myb oncogene, mediates autocrine growth and monocytic differentiation";
 RL Cell 91:185-195 (1997).
 CC -1- FUNCTION: MAY ACT AS A TRANSCRIPTION FACTOR. FOR CELL PLURIPOTENCY AND DIFFERENTIATION IN THE EMBRYO.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL; AF022151; AAB82710.1; -

DR HSSP; P14653; 1B72.
 DR TRANSFAC; T04021; -
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00024; HOMEBOX.
 DR ProDom; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00027; HOMEBOX 1; 1.
 DR PROSITE; PS00071; HOMEBOX 2; 1.
 KW Homeobox; DNA-binding; Nuclear protein; Transcription regulation.
 FT DOMAIN 31 36 POLY-PRO.
 FT DOMAIN 56 60 POLY-PRO.
 FT DOMAIN 69 80 POLY-PRO.
 FT DOMAIN 138 143 POLY-GLY.
 FT DOMAIN 240 243 POLY-ARG.
 FT DNA_BIND 239 298 HOMEBOX.
 SQ SEQUENCE 340 AA; 35855 MW; B71423225904E741 CRC64;

Query Match 18.8%; Score 299.5; DB 1; Length 340;
 Best Local Similarity 31.4%; Pred. No. 3.3e-11;
 Matches 98; Conservative 26; Mismatches 89; Indels 99; Gaps 11;

QY 1 MEKMEKSONFRIEALIAEKPPRSASPPGL-----SPGSPGPAKTDTPSPAPQA-- 53
 DB 12 MORPLGSSSTAFSISDLSIGSPPPA---PGHFVYTGYPMPVPPVLPPLPALPQAALQ 68
 QY 54 TPLGPAQFVFKPGLHLPGL-----GTLPA-----LYPPA 85
 DB 69 PLPLPA---PPPLPALPGAFCPGLAQGMALSTLMAALPSSPPAPPPPEAKRKAPPG 125
 QY 86 VY-----PLPALGGHAAFAVTAFPOLPPGAHLLKAAVAGSPLEQWTRAG----- 132
 DB 126 NFDKADLPPLPDGGGGGDDCKTGGLLPAPADAVHSLAG-----LRGPKDDPKA 179
 QY 133 -----MLVRLSD-----FHATPQALMGKS 153
 DB 180 EEEKGREENFMSDSDLDYSSDENGPAAPAREDEDCGTALEENPPSAANAAAAATGKN 239
 QY 154 RRPRTAFTSQQLLENOFLNKYLSRPRFEVATSLMTETQVKIWFQWRMKRSRK 213
 DB 240 RRRRTAFTSQQLLENOFLNKYLSRPRFEVATSLMTETQVKIWFQWRMKRSRK 299
 QY 214 AKEQGMVAVEPEK 225
 DB 300 GNASSKRGERSR 311

RESULT 3

GSH1_HUMAN STANDARD; PRT; 264 AA.
 ID GSH1_HUMAN
 AC 09H452;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Homeobox protein GSH-1.
 GN GSH1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN 11
 RP MEDLINE=21588723; PubMed=11731616;
 RA Mutsaers N., Iwasaki Y., Morishita M., Nomura A., Yamamori E., Yoshida M., Asai M., Ozaki N., Kambe F., Seo H., Oiso Y., Saito H.;
 RT "Homeobox protein Gsh-1-dependent regulation of the rat GHRH gene promoter";
 RL Mol. Endocrinol. 15:2149-2156 (2001).
 CC [2]
 RN 12
 RP SEQUENCE FROM N.A.
 RA Bates K.;
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Probable transcription factor that binds to the DNA

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: January 29, 2004, 23:05:35 ; Search time 41 Seconds
(without alignments)
1888.190 Million cell updates/sec

Title: US-09-820-598-1
Perfect score: 1589
Sequence: 1 MHRPMEKSNFRLEALAEK.....DEEEDDGMPTARKMGSVL 300

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1589	100.0	300	13 Q9YHY8	Q9YHY8 gallus gall
2	620.5	39.0	349	13 Q9YHY7	Q9YHY7 gallus gall
3	571	35.9	404	11 Q9QZM9	Q9QZM9 mus musculu
4	564.5	35.5	355	4 Q9Y648	Q9Y648 homo sapien
5	512	32.2	296	5 Q9GSP1	Q9GSP1 branchiosteo
6	487.5	30.7	237	4 Q9UDJ3	Q9UDJ3 homo sapien
7	472	29.7	261	13 Q93368	Q93368 xenopus lae
8	376	23.7	525	5 Q9VSC2	Q9VSC2 drosophila
9	288	18.1	285	11 Q8BPE6	Q8BPE6 mus musculu
10	283	17.8	255	5 Q25182	Q25182 hydra atten
11	278	17.5	255	5 Q9NEM1	Q9NEM1 hydra atten
12	273	17.2	275	5 Q96779	Q96779 branchiosteo
13	269.5	17.0	275	5 Q9UAI8	Q9UAI8 branchiosteo
14	268	16.9	718	5 Q9U718	Q9U718 chaetoptero
15	267.5	16.8	445	13 Q8AYD8	Q8AYD8 morone saxa
16	263	16.6	524	5 Q9VFC3	Q9VFC3 drosophila

17	260	16.4	214	5 Q9YIT6	Q9YIT6 nematoscell
18	253.5	16.0	316	13 Q902B2	Q902B2 brachydanio
19	252	15.9	243	13 Q9DEB1	Q9DEB1 gallus gall
20	251.5	15.8	591	5 Q966U1	Q966U1 halocynthia
21	249.5	15.7	390	13 Q93365	Q93365 brachydanio
22	248	15.6	261	5 Q17476	Q17476 hydractinia
23	247	15.5	485	5 Q24136	Q24136 drosophila
24	247	15.5	485	5 Q9V552	Q9V552 drosophila
25	246.5	15.5	281	5 Q9NB02	Q9NB02 halotile ru
26	245.5	15.4	477	5 Q9UBD7	Q9UBD7 tribolium c
27	245.5	15.4	782	5 Q97058	Q97058 drosophila
28	244.5	15.4	289	5 Q17188	Q17188 bombyx mori
29	244	15.4	280	13 Q91152	Q91152 notophthalm
30	243.5	15.3	375	13 Q9DEZ1	Q9DEZ1 xenopus lae
31	242.5	15.3	266	5 Q9GTR8	Q9GTR8 podocoryne
32	242.5	15.3	331	5 Q26469	Q26469 schistocerc
33	242.5	15.3	515	5 Q8T0H1	Q8T0H1 drosophila
34	242	15.2	240	5 Q8TE00	Q8TE00 branchiosteo
35	242	15.2	328	4 Q96CA4	Q96CA4 homo sapien
36	242	15.2	418	11 Q8CGW5	Q8CGW5 mus musculu
37	241.5	15.2	267	6 Q9GK08	Q9GK08 canis famli
38	241.5	15.2	303	13 Q9PWL2	Q9PWL2 petromyzon
39	241.5	15.2	379	5 Q9USK8	Q9USK8 schistocerc
40	241.5	15.2	410	13 Q8RWZ2	Q8RWZ2 brachydanio
41	241.5	15.2	411	5 Q9GPR6	Q9GPR6 clona intes
42	241	15.2	246	13 Q73726	Q73726 brachydanio
43	240.5	15.1	271	11 Q9R1X2	Q9R1X2 mus musculu
44	239	15.0	205	5 Q9BK18	Q9BK18 acropora ml
45	238.5	15.0	297	11 Q9QUG0	Q9QUG0 rattus norv

ALIGNMENTS

RESULT 1
ID Q9YHY8 PRELIMINARY; PRT; 300 AA.
AC Q9YHY8;
DT 01-MAY-1999 (T-REMBLrel. 10, Created)
DT 01-MAY-1999 (T-REMBLrel. 10, Last sequence update)
DT 01-MAR-2003 (T-REMBLrel. 23, Last annotation update)
DE Homeodomain protein.
GN MNR2.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98449697; PubMed=9778248;
RA Tanabe Y., William C., Jessell T.M.;
RT "Specification of Motor Neuron Identity by the MNR2 Homeodomain Protein."
RL Cell 95:67-80(1998).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL: A066860; AAC64924.1; -.
DR HSPSP: P14653; I872.
DR TRANSPAC: T03437; -.
DR InterPro: IPR001356; Homeobox.
DR Pfam: PF00046; homeobox; 1.
DR Prodom: PD000010; Homeobox; 1.
DR SMART: SM00389; HOX; 1.
DR PROSITE: PS00027; HOMEBOX_1; 1.
DR PROSITE: PS50071; HOMEBOX_2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
SQ SEQUENCE 300 AA; 32300 MW; 19F710C312EDC3B6 CRC64;
Query Match 100.0%; Score 1589; DB 13; Length 300;
Best Local Similarity 100.0%; Pred. No. 7.7e-122;
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MHRPMEKSNFRLEALAEKPPRSASPGLSPAGSPGAGRTDTPSPRAPQAAATPLGPA 60

```

Db      1 MEKMEKSONFRIBALLAEKPPRSASPPGLSPAGSGPAGRTDTPSPRAQAATPLGPAQ
      60
Qy      61 FVPRPGLHLPBGRLGTLPLALYPRAVYPLPALGGOHAFATAPQOLPPRAEHLKAVA
      120
Db      61 FVPRPGLHLPBGRLGTLPLALYPRAVYPLPALGGOHAFATAPQOLPPRAEHLKAVA
      120
Qy      121 GSFFLEQWIRAGMIVPRISDFHATPOSALMGKSRPRPTAFTSQQLLENOFKLNTKYL
      180
Db      121 GSFFLEQWIRAGMIVPRISDFHATPOSALMGKSRPRPTAFTSQQLLENOFKLNTKYL
      180
Qy      181 PKREFVATSLMTETQYKIFQNRMRKRSRKAKEQMAVEPEKPRGLGADSLPSQ
      240
Db      181 PKREFVATSLMTETQYKIFQNRMRKRSRKAKEQMAVEPEKPRGLGADSLPSQ
      240
Qy      241 PQGAGSPFVGVSGSPGTGFLCRSAELGYDDSSCSGGEDEEEDDQMDTAERKMSVL
      300
Db      241 PQGAGSPFVGVSGSPGTGFLCRSAELGYDDSSCSGGEDEEEDDQMDTAERKMSVL
      300

```

RESULT 2

```

Q9YHY7 ID Q9YHY7 PRELIMINARY; PRT; 349 AA.
AC Q9YHY7;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Homeodomain protein.
GN HB9.
OS Gallus gallus (chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
NC NCB1_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98449697; PubMed=9778248;
RA Tanabe Y., William C., Jessell T.M.;
RT "Specification of Motor Neuron Identity by the MNR2 Homeodomain
RT Protein."
RL Cell 95:67-80(1998).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; AF06861; AAC4925.1; -.
DR HSSP; P14653; 1B72.
DR TRANSFAC; T03436; -.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SMO0389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX 1; 1.
DR PROSITE; PS00071; HOMEBOX 2; 1.
DR DNA-binding; Homeobox; Nuclear protein.
SQ SEQUENCE 349 AA; 37013 MW; C1CF300EP814BBA8 CRC64;

```

Query Match 39.0%; Score 620.5; DB 13; Length 349;
 Best Local Similarity 47.3%; Pred. No. 1e-42; Indels 14;
 Matches 159; Conservative 30; Mismatches 66; Gaps 79; Gaps 14;

```

Qy      5 MEKSONFRIBALLAEKPPRSASPPGLSPAGSGPAGRTDTPSPRAQAATPLGPAQ
      39
Db      1 MEKSONFRIBALLAEKPPRSASPPGLSPAGSGPAGRTDTPSPRAQAATPLGPAQ
      60
Qy      40 GRITTPB-----RAPQA--TPUGPAGVPRKGL--LHLPGRLGTLPLALYPRAV
      86
Db      61 PRTPSPRPRLPLPACALPFAALFGGCGGPGGHONHALGHPAGPG--GPGLYGHV
      117
Qy      87 YPLPAGGHAFAVY-----AFPGQRPAGAHKAAVAGSPLEQWIR--AGMLVPR
      137
Db      118 YGYPALGQHPALSYSYQVGAHP--AHPSADPTKLS--AGTFQDDQLRASTGMLPK
      174
Qy      138 LSHDHAQPOSALMGKSRPRPTAFTSQQLLENOFKLNTKYLSPKREFVATSLMTETQV
      197

```

```

Db      175 MPDEGSAQSNLKGCRPRPTAFTSQQLLEHOFKLNTKYLSPKREFVATSLMTETQV
      234
Qy      198 KIFQNRMRKRSRKAKEQMAVEPEKPRGLGAD-----ESLPSQAGSDSPF
      251
Db      235 KIFQNRMRKRSRKAKEQMAVEPEKPRGLGAD-----ESLPSQAGSDSPF
      288
Qy      252 VGSPTGFLCRSAELGYDDSSCSGGEDEEEDD 287
Db      289 -----RLREL---PDSEPEDEEEEEE 310

```

RESULT 3

```

Q9QZM9 ID Q9QZM9 PRELIMINARY; PRT; 404 AA.
AC Q9QZM9;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Homeodomain protein HB9.
GN HLB9.
OS Mus musculus (mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NC NCB1_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Swiss Webster/NTI;
RX MEDLINE=99400556; PubMed=10471502;
RA Harrison K.A., Thaler J., Pfaff S.L., Gu H., Kehrl J.H.;
RT "Pancreas dorsal lobe agenesis and abnormal islets of Langerhans in
RT Hlb9-deficient mice."
RL Nat. Genet. 23:71-75(1999).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; AF153046; AAD49613.1; -.
DR HSSP; P14653; 1B72.
DR TRANSFAC; T03421; -.
DR MGD; MGI:109160; HLB9.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SMO0389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX 1; 1.
DR PROSITE; PS00071; HOMEBOX 2; 1.
DR DNA-binding; Homeobox; Nuclear protein.
SQ SEQUENCE 404 AA; 41330 MW; C25D8DBB40CFB29 CRC64;

```

Query Match 35.9%; Score 571; DB 11; Length 404;
 Best Local Similarity 40.6%; Pred. No. 1.4e-38;
 Matches 159; Conservative 29; Mismatches 70; Indels 134; Gaps 17;

```

Qy      5 MEKSONFRIBALLAEKPPRSASPPGLSPAGSGPAGRTDTPSPRAQAATPLGPAQ
      31
Db      1 MEKSONFRIBALLAEKPPRSASPPGLSPAGSGPAGRTDTPSPRAQAATPLGPAQ
      60
Qy      32 PAGES-----PGAPRTTPSPRAPQAATPLGPAQVPRKGL-----HL
      70
Db      61 PASSEATTAEDRLRAESPRLAA-----HCALLKPRGLGAGGGAAGGPTNNH
      116
Qy      71 PGP-----IGTLP-----ALYPRVY-----PLPALG
      93
Db      117 AHFGAALAAAAAAGGALGHPGAAGAGLPAQAALYGHVYSAAAAAALA
      176
Qy      94 GQHAAPVYAFPLQ-----PPGAHLLKAAVAGSPLEQWIR--AGMLVPRISDFHATPQ
      146
Db      177 GQHPALVY--SYPOVGAHPHAPADPIKLG--ASTQDDQMLRASTAGMILPMPDPSQQAQ
      234
Qy      147 SALMGKSRPRPTAFTSQQLLENOFKLNTKYLSPKREFVATSLMTETQYKIFQNR
      206
Db      235 SNLIGKCRPRPTAFTSQQLLEHOFKLNTKYLSPKREFVATSLMTETQYKIFQNR
      294
Qy      207 KMKSRKAKEQMAVEPEKPRGLG-----KADESLLPSQAGSDSPFVGCSP
      256

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - protein search, using SW model

Run on: January 29, 2004, 23:06:55 ; Search time 21 Seconds
(without alignments)

604.441 Million cell updates/sec

Title: US-09-820-598-1

Sequence: 1 MHKPMKXQNFRIEALLAEK.....DEEEEDDMGTAEKMGSVL 300

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database:

Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PTUG_COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1589	100.0	300	4	US-09-162-524-1
2	620.5	39.0	349	4	US-09-162-524-3
3	266.5	16.8	283	2	US-08-202-044-2
4	266.5	16.8	283	3	US-08-751-344B-2
5	266	16.7	284	3	US-08-589-028-6
6	266	16.7	284	3	US-08-784-582-6
7	266	16.7	284	3	US-08-785-271-6
8	265.5	16.7	283	1	US-08-583-672-2
9	254	16.0	284	2	US-08-320-148B-2
10	254	16.0	284	3	US-08-031-898-2
11	250.5	15.8	301	3	US-09-095-117-6
12	250.5	15.8	301	3	US-09-095-117-8
13	232	14.6	335	3	US-09-095-117-2
14	232	14.6	335	3	US-09-095-117-4
15	229	14.4	303	2	US-08-203-532F-2
16	229	14.4	303	3	US-09-078-465-2
17	229	14.4	303	5	PCT-US95-01882A-2
18	228.5	14.4	302	2	US-08-203-532F-4
19	228.5	14.4	302	3	US-09-950-860-16
20	228.5	14.4	302	3	US-09-078-465-4
21	228.5	14.4	302	5	PCT-US95-01882A-4
22	224	14.1	367	4	US-09-009-816-2
23	218	13.7	333	2	US-08-712-948-1
24	217	13.7	330	2	US-08-712-948-2
25	197.5	12.4	147	4	US-09-439-313-336
26	197.5	12.4	147	4	US-09-352-616A-336
27	197.5	12.4	147	4	US-09-232-149A-336

28	197.5	12.4	240	4	US-09-636-735A-2	Sequence 2, Appl
29	197.5	12.4	240	4	US-09-636-735A-12	Sequence 12, Appl
30	197	12.4	61	2	US-08-757-316C-30	Sequence 30, Appl
31	197	12.4	61	3	US-08-751-344B-9	Sequence 9, Appl
32	196	12.3	60	3	US-08-751-344B-5	Sequence 5, Appl
33	194	12.2	61	3	US-08-751-344B-7	Sequence 7, Appl
34	192.5	12.1	371	2	US-08-442-809A-76	Sequence 76, Appl
35	188	11.8	61	3	US-08-751-344B-8	Sequence 8, Appl
36	184	11.6	53	2	US-08-549-004A-11	Sequence 11, Appl
37	184	11.6	53	3	US-09-051-982A-11	Sequence 11, Appl
38	184	11.6	61	2	US-08-202-044-3	Sequence 3, Appl
39	184	11.6	61	3	US-08-751-344B-3	Sequence 3, Appl
40	184	11.6	61	3	US-08-751-344B-6	Sequence 6, Appl
41	183	11.5	53	2	US-08-549-004A-12	Sequence 12, Appl
42	183	11.5	53	3	US-09-051-982A-12	Sequence 12, Appl
43	182.5	11.5	302	4	US-08-957-351-7	Sequence 7, Appl
44	182.5	11.5	302	4	US-08-957-351-30	Sequence 30, Appl
45	182.5	11.5	313	4	US-08-957-351-9	Sequence 9, Appl

ALIGNMENTS

```

RESULT 1
US-09-162-524-1
; Sequence 1, Application US/09162524
; Patent No. 6387656
; GENERAL INFORMATION:
; APPLICANT: Janssen, Thomas M.
; APPLICANT: Tanabe, Yasuo
; TITLE OF INVENTION: Gene Encoding MNR2 and Uses Thereof
; FILE REFERENCE: 57477/jbw/w1
; CURRENT APPLICATION NUMBER: US/09/162,524
; CURRENT FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 300
; TYPE: PRT
; ORGANISM: chick embryo
US-09-162-524-1

Query Match      100.0%; Score 1589; DB 4; Length 300;
Best Local Similarity 100.0%; Pred. No. 1.9e-139;
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MHKPMKXQNFRIEALLAEKPPRSASPPGGLSPAGSPGAGRTDTPSPRAPOAATPLGPAAG 60
        |||
DB      1 MHKPMKXQNFRIEALLAEKPPRSASPPGGLSPAGSPGAGRTDTPSPRAPOAATPLGPAAG 60

QY      61 FVVRKPGILHPSGGLGTLPLPYPPAYVPLPALGGQHAFAFYTAFPOLPPGAEHLKAVA 120
        |||
DB      61 FVVRKPGILHPSGGLGTLPLPYPPAYVPLPALGGQHAFAFYTAFPOLPPGAEHLKAVA 120

QY      121 GSPPLEOWIRAGMLVPRLSFPHATPOSALMGKRRPRPTATTSQOLELENOFKLNTKYSR 180
        |||
DB      121 GSPPLEOWIRAGMLVPRLSFPHATPOSALMGKRRPRPTATTSQOLELENOFKLNTKYSR 180

QY      181 PKRFEVATSLMTETQYKIWFQNRMMKRSRAKEQMAVEBEKPRGLGKADESLIPSO 240
        |||
DB      181 PKRFEVATSLMTETQYKIWFQNRMMKRSRAKEQMAVEBEKPRGLGKADESLIPSO 240

QY      241 PQOAGDSPPFVCGSPRTGFLCSAELGYDPDSSCGSEDEEDDMGTAEKMGSVL 300
        |||
DB      241 PQOAGDSPPFVCGSPRTGFLCSAELGYDPDSSCGSEDEEDDMGTAEKMGSVL 300

RESULT 2
US-09-162-524-3
; Sequence 3, Application US/09162524
; Patent No. 6387656
; GENERAL INFORMATION:

```

APPLICANT: Jessell, Thomas M.
 APPLICANT: Tanabe, Yasuo
 APPLICANT: Williams, Christopher
 TITLE OF INVENTION: Gene Encoding MMR2 and Uses Thereof
 FILE REFERENCE: 57477/jpw/w1
 CURRENT APPLICATION NUMBER: US/09/162,524
 CURRENT FILING DATE: 1998-09-29
 NUMBER OF SEQ ID NOS: 4
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO: 3
 LENGTH: 349
 TYPE: PRT
 ORGANISM: chick embryo
 US-09-162-524-3

Query Match 39.0%; Score 620.5; DB 4; Length 349;
 Best Local Similarity 47.3%; Pred. No. 1.4e-49;
 Matches 159; Conservative 30; Mismatches 68; Indels 79; Gaps 14;

QY 5 MEKSONRFEALAEKPPRSA-----SPPGLSPAGS-----GPA 39
 DB 1 MEKSKNFRIDALAVDPKAAAGAPLALVTGSGGSGSPSSSSSSSSSELPA DC 60
 QY 40 GRTDTPSP-----RAPQA--TPLGPAFVPRKGL--LHLPGGLGTLPALYPPAV 86
 DB 61 PRDPSPPRLRLPAHCAFLPKAFLGGGPGGHPQHHLGLHPAGP--GPGLYGHPV 117
 QY 87 YPLPALGGHAAFYT-----APQLPPGAELHKAIVAGSFLBQWIR--AGMLVPR 137
 DB 118 YGVYALGGHPLALSYSSVOVGAPH--AHPSADPIKLS--AGTFQIDWLRASTAGMLPK 174
 QY 138 LSDHATPOSALMGSRPRTAFTSQQLLENOFKLKYSRKRPVATSLMLTETOV 197
 DB 175 MPDGGSAQSULKCRPRAFTSQQLLEHOFKLKYSRKREVAISLMLTETOV 234
 QY 198 KIMFONRMKWKSRKAKKEQMAVEPEKPRGLGKAD-----ESLLPSQPGQAGDSPEF 251
 DB 235 KIMFONRMKWKSRKAKKEQA--AQEAENEKGGGGGSDSGPRELLLPGEKGGGR----- 288
 QY 252 VGCSPGTGFLCRSAELGYDPSSCGGHEDEEEDD 287
 DB 289 -----RLREL---PDSEPEDEEEEEEEEEE 310

RESULT 3
 US-08-202-044-2
 Sequence 2, Application US/08202044
 Patent No. 5858973
 GENERAL INFORMATION:
 APPLICANT: Habener M.D., Joel F.
 APPLICANT: Miller Ph.D., Christopher P.
 TITLE OF INVENTION: NOVEL TRANSCRIPTION FACTOR AND USES
 NUMBER OF SEQUENCES: 29
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Meingarten, Schurgin, Gagnebin & Hayes
 STREET: Ten Post Office Square
 CITY: Boston
 STATE: MA
 COUNTRY: US
 ZIP: 02109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/202,044
 FILING DATE: 23-FEB-1994
 CLASSIFICATION: 800
 ATTORNEY/AGENT INFORMATION:
 NAME: Williams Ph.D., Kathleen A.
 REGISTRATION NUMBER: 34,380

REFERENCE/DOCKET NUMBER: MGH-124XX
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 542-2290
 TELEFAX: (617) 451-0313
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 283 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-202-044-2

Query Match 16.8%; Score 266.5; DB 2; Length 283;
 Best Local Similarity 32.9%; Pred. No. 6.7e-17;
 Matches 92; Conservative 23; Mismatches 106; Indels 59; Gaps 13;

QY 21 PPSASPP-----GLSPAGSPG--AGRTYT-----PSPAQAATPLG--PAG----- 60
 DB 26 PPSANPPACLYMGKPPPTPOFAGSLGTLBQSGPPDISPYEVPLADDPAGAHLLHH 85
 QY 61 FVPRPGLHLPGGLGTLPALYPPAVYPLPALGGHAAFAVTAAPOLPPGAELHKAIVA 120
 DB 86 LPAQLGLAH--PPPG-----PFPN--CTETGGLBEPKRVHLPPFMKSTKAHAM 130
 QY 121 GSPFLEQWIRAGMLVPRISDFHATPOSALMGSRPRTAFTSQQLLENOFKLKYLSR 180
 DB 131 KS-----QW--AG-----GAYAAPE-----ENKRTTATYTRQLLELEKEFLFNKYISR 173
 QY 181 PKREVAISLMLTETOVKIMFONRMKWKSRKAKKEQ-----GNAVEPEKRGGLKAD 233
 DB 174 PRVEELAVMLNLTERRHKKIMFONRMKWKKEEDKSSGTTSGGGGGGEEPEQCAVTSGE 233
 QY 234 ESLLPSQPGQAGDSPEFVGCSPGTGFLCRSAELGYDPDS 273
 DB 234 ELIALPKPPPGGVPSGVPAABEGRLPSGLSASPOPSS 273

RESULT 4
 US-08-751-344B-2
 Sequence 2, Application US/08751344B
 Patent No. 6210960
 GENERAL INFORMATION:
 APPLICANT: Habener M.D., Joel F.
 APPLICANT: Miller Ph.D., Christopher P.
 TITLE OF INVENTION: NOVEL TRANSCRIPTION FACTOR AND USES
 NUMBER OF SEQUENCES: 29
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Banner & Witcoff, Ltd.
 STREET: One Financial Center
 CITY: Boston
 STATE: MA
 COUNTRY: US
 ZIP: 02111
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: WordPerfect 6.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/751,344B
 FILING DATE: 19-NOV-6210960-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/202,044
 FILING DATE: 23-FEB-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Williams Ph.D., Kathleen M.
 REGISTRATION NUMBER: 34,380
 REFERENCE/DOCKET NUMBER: 96,137-A (11274/02148)
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 345-9100
 TELEFAX: (617) 345-9111
 INFORMATION FOR SEQ ID NO: 2:

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 23:09:16 ; Search time 34 Seconds
(without alignments)
1833.742 Million cell updates/sec

Title: US-09-820-598-1

Perfect score: 1589
Sequence: 1 MHKPMKSONFRLEALAEK.....DEBEEDDGMTARKKGSVL 300

Scoring table: BLOSUM62
Gap 10.0 , Gapext 0.5

Searched: 789580 seqs, 207824079 residues

Total number of hits satisfying chosen parameters: 789580

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubppaa/PCR_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubppaa/PCRUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1589	100.0	300	11	US-09-820-598-1
2	1589	100.0	300	14	US-10-095-932-1
3	620.5	39.0	349	11	US-09-820-598-3
4	620.5	39.0	349	14	US-10-095-932-3
5	274	17.2	283	12	US-10-162-952-1
6	266.5	16.8	416	15	US-10-305-667A-2
7	265	16.7	443	15	US-10-305-667A-15
8	259	16.3	356	15	US-10-177-293-214
9	254	16.0	284	9	US-09-759-847-2
10	250.5	15.8	301	9	US-09-850-260-6
11	250.5	15.8	301	10	US-09-850-258-6
12	250.5	15.8	304	9	US-09-850-260-8
13	250.5	15.8	304	10	US-09-850-258-8
14	241.5	15.2	297	12	US-10-295-027-630
15	241	15.2	252	12	US-10-094-749-1844

16	232	14.6	330	15	US-10-157-031-70	Sequence 70, Appl
17	232	14.6	335	9	US-09-850-260-2	Sequence 2, Appl1
18	232	14.6	335	9	US-09-850-260-4	Sequence 4, Appl1
19	232	14.6	335	10	US-09-850-258-2	Sequence 2, Appl1
20	232	14.6	335	10	US-09-850-258-4	Sequence 4, Appl1
21	232	14.6	335	12	US-10-116-275-181	Sequence 181, App
22	229	14.4	303	12	US-09-940-673-2	Sequence 2, Appl1
23	228.5	14.4	302	12	US-09-940-673-4	Sequence 4, Appl1
24	221	13.9	217	15	US-10-097-340-129	Sequence 129, App
25	221	13.9	254	15	US-10-012-456A-54	Sequence 54, Appl
26	221	13.9	311	15	US-10-012-456A-38	Sequence 38, Appl
27	219.5	13.8	324	10	US-09-749-728B-9	Sequence 9, Appl1
28	219.5	13.8	324	12	US-10-177-390-26	Sequence 26, Appl
29	216	13.6	265	9	US-09-819-254-2	Sequence 2592, Ap
30	214	13.5	265	10	US-09-998-598-2592	Sequence 1184, Ap
31	214	13.5	265	12	US-10-295-027-1184	Sequence 281, App
32	210	13.2	407	15	US-10-157-031-281	Sequence 2, Appl1
33	209	13.2	234	11	US-09-105-470-2	Sequence 2, Appl1
34	209	13.2	234	12	US-10-341-434-63	Sequence 63, Appl
35	209	13.2	234	12	US-10-614-275-2	Sequence 2, Appl1
36	209	13.2	234	15	US-10-205-823-283	Sequence 283, App
37	209	13.2	241	9	US-09-759-143-842	Sequence 842, App
38	209	13.2	241	9	US-09-780-669-842	Sequence 842, App
39	209	13.2	241	10	US-09-822-827-842	Sequence 842, App
40	209	13.2	241	10	US-09-895-793-842	Sequence 842, App
41	209	13.2	241	10	US-09-895-814-842	Sequence 842, App
42	209	13.2	241	12	US-10-144-678A-842	Sequence 842, App
43	209	13.2	241	12	US-10-294-025-842	Sequence 842, App
44	209	13.2	241	12	US-10-012-895-842	Sequence 842, App
45	208	13.1	269	12	US-10-116-275-150	Sequence 190, App

ALIGNMENTS

RESULT 1
US-09-820-598-1
; Publication 1, Application US/09820598
; Publication No. US20030104374A1
; GENERAL INFORMATION:
; APPLICANT: Thomas M. Jessell et al.
; TITLE OF INVENTION: GENE ENCODING MNF2 AND USES THEREOF
; FILE REFERENCE: 0575/57477-A-PCT-US/JPM/SHS/MVM
; CURRENT APPLICATION NUMBER: US/09/820,598
; CURRENT FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 300
; TYPE: PRT
; ORGANISM: chick embryo
US-09-820-598-1

Query Match 100.0%; Score 1589; DB 11; Length 300;
Best Local Similarity 100.0%; Pred. No. 9.4e-110;
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MHKPMKSONFRLEALAEKPPRSASPPGSLSPGSGPPACRTTTPSPRAQAATPLGPAG	60
DB	1	MHKPMKSONFRLEALAEKPPRSASPPGSLSPGSGPPACRTTTPSPRAQAATPLGPAG	60
QY	61	FVPRKPGILHPPGGLGTLPLALYPPAYYPLPALGGGAAAFATFAPOLPPGAGHLKAAYA	120
DB	61	FVPRKPGILHPPGGLGTLPLALYPPAYYPLPALGGGAAAFATFAPOLPPGAGHLKAAYA	120
QY	121	GSFPLEQWIRAGLTVPRLSDFHATPOSALMGKSRPRRTAFTSQQLLELENOFLNKYLSR	180
DB	121	GSFPLEQWIRAGLTVPRLSDFHATPOSALMGKSRPRRTAFTSQQLLELENOFLNKYLSR	180
QY	181	PKRFVATSLMLTETQYKTFQNRKMKWSRRAKQGAHVAEPEKRGJGAKADESLIPSG	240
DB	181	PKRFVATSLMLTETQYKTFQNRKMKWSRRAKQGAHVAEPEKRGJGAKADESLIPSG	240

Qy 241 POGAGDSPEFVGCSPTGFLCRSAELGYDDSSCGGGEDEEEDDGMDTAERKMGSVL 300
 Db 241 POGAGDSPEFVGCSPTGFLCRSAELGYDDSSCGGGEDEEEDDGMDTAERKMGSVL 300

RESULT 2
 US-10-095-932-1
 ; Sequence 1, Application US/10095932
 ; Publication No. US20020197678A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Jesell, Thomas M.
 ; APPLICANT: Tanabe, Yasuo
 ; APPLICANT: William, Christopher
 ; TITLE OF INVENTION: Gene Encoding MNR2 and Uses Thereof
 ; FILE REFERENCE: 57477/jpw/w1
 ; CURRENT APPLICATION NUMBER: US/10/095,932
 ; CURRENT FILING DATE: 2002-03-11
 ; PRIOR APPLICATION NUMBER: US/09/162,524
 ; PRIOR FILING DATE: 1998-09-29
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: Patent In Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 300
 ; TYPE: PRT
 ; ORGANISM: chick embryo
 US-10-095-932-1

Query Match 100.0%; Score 1589; DB 14; Length 300;
 Best Local Similarity 100.0%; Pred. No. 9,4e-110;
 Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MHKMEKSONFRIALIAEKPPRSASPPGLSPAGSPGARTDTPSPRAPOATPLGPAG 60
 Db 1 MHKMEKSONFRIALIAEKPPRSASPPGLSPAGSPGARTDTPSPRAPOATPLGPAG 60
 Qy 61 FVPRGLHLHPGPGIGTLPALYPAYVPLPALGGQHAFAVTAEPOLPPGAEHLKAVA 120
 Db 61 FVPRGLHLHPGPGIGTLPALYPAYVPLPALGGQHAFAVTAEPOLPPGAEHLKAVA 120
 Qy 121 GSPFLEOMIRAGMIVPRISDFHATPQSLMKSKRRPRATFSQQLLENOFKLNTLSR 180
 Db 121 GSPFLEOMIRAGMIVPRISDFHATPQSLMKSKRRPRATFSQQLLENOFKLNTLSR 180
 Qy 181 PKREVATSLMLTETQYKIFQNRMKKRSKAKBOGMAVEPKPGGLGKADESLPSQ 240
 Db 181 PKREVATSLMLTETQYKIFQNRMKKRSKAKBOGMAVEPKPGGLGKADESLPSQ 240
 Qy 241 POGAGDSPEFVGCSPTGFLCRSAELGYDDSSCGGGEDEEEDDGMDTAERKMGSVL 300
 Db 241 POGAGDSPEFVGCSPTGFLCRSAELGYDDSSCGGGEDEEEDDGMDTAERKMGSVL 300

RESULT 3
 US-09-820-598-3
 ; Sequence 3, Application US/09820598
 ; Publication No. US20030104374A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Thomas M. Jesell et al.
 ; TITLE OF INVENTION: GENE ENCODING MNR2 AND USES THEREOF
 ; FILE REFERENCE: 0575/57477-A-PCT-US/jpw/shs/wm
 ; CURRENT APPLICATION NUMBER: US/09/820,598
 ; CURRENT FILING DATE: 2001-03-29
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: Patent In version 3.1
 ; SEQ ID NO 3
 ; LENGTH: 349
 ; TYPE: PRT
 ; ORGANISM: chick embryo
 US-09-820-598-3

Query Match 39.0%; Score 620.5; DB 11; Length 349;
 Best Local Similarity 47.3%; Pred. No. 4.5e-38;
 Matches 159; Conservative 30; Mismatches 68; Indels 79; Gaps 14;

Qy 5 MEKSONFRIALIAEKPPRSA-----SPPGLSPAGS-----PGPA 39
 Db 1 MEKSONFRIALIAEKPPRSA-----SPPGLSPAGS-----PGPA 39
 Qy 40 GRDTTSP-----RAPQAA--TPLPAGVVPKGL--LHPGGLGTLPALYPAY 86
 Db 40 GRDTTSP-----RAPQAA--TPLPAGVVPKGL--LHPGGLGTLPALYPAY 86
 Qy 61 PRDTSPPRLLPAHCHALLPKAFLGGGGGPGGHPOHNLGLHPAGG--GGLYGHV 117
 Db 61 PRDTSPPRLLPAHCHALLPKAFLGGGGGPGGHPOHNLGLHPAGG--GGLYGHV 117
 Qy 87 YPLPAGGQHAFAVY-----APPOLPPGAEHLKAAVAGSPFLGQWIR---AGMLVPR 137
 Db 87 YPLPAGGQHAFAVY-----APPOLPPGAEHLKAAVAGSPFLGQWIR---AGMLVPR 137
 Qy 118 YGYPALGGQHPALSYSTSYQVGAHP--AHPSADPIKLS--AGTQLODMWLASTRAGMILPK 174
 Db 118 YGYPALGGQHPALSYSTSYQVGAHP--AHPSADPIKLS--AGTQLODMWLASTRAGMILPK 174
 Qy 138 LSPFHATPQSLMGSKRRPRATFTSQQLLENOFKLNTLSRPRKFEVATSLMTETQV 197
 Db 138 LSPFHATPQSLMGSKRRPRATFTSQQLLENOFKLNTLSRPRKFEVATSLMTETQV 197
 Qy 175 MPDFGSAQGNLGLCKRRPRATFTSQQLLENOFKLNTLSRPRKFEVATSLMTETQV 234
 Db 175 MPDFGSAQGNLGLCKRRPRATFTSQQLLENOFKLNTLSRPRKFEVATSLMTETQV 234
 Qy 198 KIFQNRMKKRSKRAKEQMAVEPKPGGLGKAD-----ESLLPSQOGAGDSPEF 251
 Db 198 KIFQNRMKKRSKRAKEQMAVEPKPGGLGKAD-----ESLLPSQOGAGDSPEF 251
 Qy 235 KIFQNRMKKRSKRAKEQMAVEPKPGGLGKAD-----ESLLPSQOGAGDSPEF 288
 Db 235 KIFQNRMKKRSKRAKEQMAVEPKPGGLGKAD-----ESLLPSQOGAGDSPEF 288
 Qy 252 VCGSPGTGFLCRSAELGYDDSSCGGGEDEEEDD 287
 Db 252 VCGSPGTGFLCRSAELGYDDSSCGGGEDEEEDD 287
 Qy 289 -----RULEL---PDSEPEDEEEEEE 310
 Db 289 -----RULEL---PDSEPEDEEEEEE 310

RESULT 4
 US-10-095-932-3
 ; Sequence 3, Application US/10095932
 ; Publication No. US20020197678A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Jesell, Thomas M.
 ; APPLICANT: Tanabe, Yasuo
 ; APPLICANT: William, Christopher
 ; TITLE OF INVENTION: Gene Encoding MNR2 and Uses Thereof
 ; FILE REFERENCE: 57477/jpw/w1
 ; CURRENT APPLICATION NUMBER: US/10/095,932
 ; CURRENT FILING DATE: 2002-03-11
 ; PRIOR APPLICATION NUMBER: US/09/162,524
 ; PRIOR FILING DATE: 1998-09-29
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: Patent In Ver. 2.0
 ; SEQ ID NO 3
 ; LENGTH: 349
 ; TYPE: PRT
 ; ORGANISM: chick embryo
 US-10-095-932-3

Query Match 39.0%; Score 620.5; DB 14; Length 349;
 Best Local Similarity 47.3%; Pred. No. 4.5e-38;
 Matches 159; Conservative 30; Mismatches 68; Indels 79; Gaps 14;

Qy 5 MEKSONFRIALIAEKPPRSA-----SPPGLSPAGS-----PGPA 39
 Db 1 MEKSONFRIALIAEKPPRSA-----SPPGLSPAGS-----PGPA 39
 Qy 40 GRDTTSP-----RAPQAA--TPLPAGVVPKGL--LHPGGLGTLPALYPAY 86
 Db 40 GRDTTSP-----RAPQAA--TPLPAGVVPKGL--LHPGGLGTLPALYPAY 86
 Qy 61 PRDTSPPRLLPAHCHALLPKAFLGGGGGPGGHPOHNLGLHPAGG--GGLYGHV 117
 Db 61 PRDTSPPRLLPAHCHALLPKAFLGGGGGPGGHPOHNLGLHPAGG--GGLYGHV 117
 Qy 87 YPLPAGGQHAFAVY-----APPOLPPGAEHLKAAVAGSPFLGQWIR---AGMLVPR 137
 Db 87 YPLPAGGQHAFAVY-----APPOLPPGAEHLKAAVAGSPFLGQWIR---AGMLVPR 137
 Qy 118 YGYPALGGQHPALSYSTSYQVGAHP--AHPSADPIKLS--AGTQLODMWLASTRAGMILPK 174
 Db 118 YGYPALGGQHPALSYSTSYQVGAHP--AHPSADPIKLS--AGTQLODMWLASTRAGMILPK 174
 Qy 138 LSPFHATPQSLMGSKRRPRATFTSQQLLENOFKLNTLSRPRKFEVATSLMTETQV 197
 Db 138 LSPFHATPQSLMGSKRRPRATFTSQQLLENOFKLNTLSRPRKFEVATSLMTETQV 197
 Qy 175 MPDFGSAQGNLGLCKRRPRATFTSQQLLENOFKLNTLSRPRKFEVATSLMTETQV 234
 Db 175 MPDFGSAQGNLGLCKRRPRATFTSQQLLENOFKLNTLSRPRKFEVATSLMTETQV 234
 Qy 198 KIFQNRMKKRSKRAKEQMAVEPKPGGLGKAD-----ESLLPSQOGAGDSPEF 251
 Db 198 KIFQNRMKKRSKRAKEQMAVEPKPGGLGKAD-----ESLLPSQOGAGDSPEF 251
 Qy 235 KIFQNRMKKRSKRAKEQMAVEPKPGGLGKAD-----ESLLPSQOGAGDSPEF 288
 Db 235 KIFQNRMKKRSKRAKEQMAVEPKPGGLGKAD-----ESLLPSQOGAGDSPEF 288
 Qy 252 VCGSPGTGFLCRSAELGYDDSSCGGGEDEEEDD 287
 Db 252 VCGSPGTGFLCRSAELGYDDSSCGGGEDEEEDD 287